



IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

44 7-7-02  
Atty. Blumenthal 09/744/0305  
*Priority Papers*

Applicant: Atsushi MURASHIMA  
Title: SPEECH SIGNAL DECODING  
METHOD AND APPARATUS  
Appl. No.: 09/627,421  
Filing Date: 07/27/2000  
Examiner: Unknown  
Art Unit: 2741

**CLAIM FOR CONVENTION PRIORITY**

Commissioner for Patents  
Washington, D.C. 20231

Sir:

The benefit of the filing date of the following prior foreign application filed in the following foreign country is hereby requested, and the right of priority provided in 35 U.S.C. § 119 is hereby claimed.

In support of this claim, filed herewith is a certified copy of said original foreign application:

- JAPAN Patent Application No. 11-214292 filed 07/28/1999.

Respectfully submitted,

Date June 21, 2002

By

*D. A. Blumenthal*  
# 44,740

FOLEY & LARDNER  
Customer Number: 22428



22428

PATENT TRADEMARK OFFICE  
Telephone: (202) 672-5407  
Facsimile: (202) 672-5399

David A. Blumenthal  
Attorney for Applicant  
Registration No. 26,257



日 本 国 特 許 庁  
PATENT OFFICE  
JAPANESE GOVERNMENT

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出 願 年 月 日

Date of Application:

1999年 7月28日

出 願 番 号

Application Number:

平成11年特許願第214292号

出 願 人

Applicant (s):

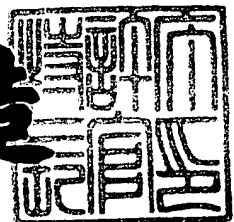
日本電気株式会社

CERTIFIED COPY OF  
PRIORITY DOCUMENT

2000年 9月 1日

特許庁長官  
Commissioner,  
Patent Office

及 川 耕 造



出証番号 出証特2000-3070524

【書類名】 特許願  
 【整理番号】 33509585  
 【あて先】 特許庁長官殿  
 【国際特許分類】 H04B 14/04  
 H04B 15/00  
 G10L 9/14

【発明者】

【住所又は居所】 東京都港区芝五丁目 7 番 1 号  
 本電気株式会社内

日

【氏名】 村島 淳

【特許出願人】

【識別番号】 000004237

【氏名又は名称】 日本電気株式会社

【代理人】

【識別番号】 100082935

【弁理士】

【氏名又は名称】 京本 直樹

【電話番号】 03-3454-1111

【選任した代理人】

【識別番号】 100082924

【弁理士】

【氏名又は名称】 福田 修一

【電話番号】 03-3454-1111

【選任した代理人】

【識別番号】 100085268

【弁理士】

【氏名又は名称】 河合 信明

【電話番号】 03-3454-1111

【手数料の表示】

【予納台帳番号】 008279

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【包括委任状番号】 9115699

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 音声信号復号方法及び音声信号復号装置及び音声信号符号化復号方法及び音声信号符号化復号装置

【特許請求の範囲】

【請求項 1】 受信した信号から少なくとも音源信号とゲインとフィルタ係数の情報を復号し、前記音源信号に前記ゲインを乗じて得られる励振信号により前記フィルタ係数で構成するフィルタを駆動することによって音声信号を復号する音声信号復号方法において、前記復号した情報を用いて前記受信した信号について有音部と無音部との識別を行ない、前記無音部にて前記情報のうち前記ゲインと前記フィルタ係数の少なくとも一つに対して前記情報に基づく平滑化処理を施した結果を用いて前記音声信号の復号を行うことを特徴とする音声信号復号方法。

【請求項 2】 受信した信号から少なくとも音源信号とゲインとフィルタ係数の情報を復号し、前記音源信号に前記ゲインを乗じて得られる励振信号により前記フィルタ係数で構成するフィルタを駆動することによって音声信号を復号する音声信号復号方法において、前記復号した情報を用いて前記受信した信号について有音部と無音部との識別を行うとともに前記復号した情報により無音部を分類し、前記無音部にて前記情報のうち前記ゲインと前記フィルタ係数の少なくとも一つに対して前記分類に応じて平滑化処理を施した結果を用いて前記音声信号の復号を行うことを特徴とする音声信号復号方法。

【請求項 3】 前記フィルタ係数とその長時間平均との差分に基づく変動量を長時間平均した値を用いて前記識別と前記分類の少なくとも一つを行なうことを特徴とする請求項 1 または 2 記載の音声信号復号方法。

【請求項 4】 前記受信した信号から入力信号のピッチ周期性和パワーの情報を復号し、前記識別及び前記分類の少なくとも一つにて前記ピッチ周期性和前記パワーの情報の少なくとも一つを用いることを特徴とする請求項 1、2 または 3 記載の音声信号復号方法。

【請求項 5】 復号した前記励振信号と前記音声信号から入力信号のピッチ周期性和パワーの情報を推定し、前記識別及び前記分類の少なくとも一つにて

前記推定したピッチ周期性と前記推定したパワーの情報の少なくとも一つを用いることを特徴とする請求項 1、2 または 3 記載の音声信号復号方法。

【請求項 6】 前記復号した情報のうち前記フィルタ係数の情報から得られる値を予め定められた閾値と比較することにより前記無音部を分類することを特徴とする請求項 2 記載の音声信号復号方法。

【請求項 7】 受信した信号から少なくとも音源信号とゲインとフィルタ係数の情報を復号し、前記音源信号に前記ゲインを乗じて得られる励振信号により前記フィルタ係数で構成するフィルタを駆動することによって音声信号を復号する音声信号復号装置において、前記復号した情報を用いて前記受信した信号について有音部と無音部との識別を行なう有音/無音識別回路と、前記有音/無音識別回路で識別された前記無音部にて前記情報のうち前記ゲインと前記フィルタ係数の少なくとも一つに対して前記情報に基づく平滑化処理を施す平滑化回路と、前記平滑化回路によって平滑化処理を施した結果を用いて前記音声信号の復号を行う復号回路を含んで構成されることを特徴とする音声信号復号装置。

【請求項 8】 受信した信号から少なくとも音源信号とゲインとフィルタ係数の情報を復号し、前記音源信号に前記ゲインを乗じて得られる励振信号により前記フィルタ係数で構成するフィルタを駆動することによって音声信号を復号する音声信号復号装置において、前記復号した情報を用いて前記受信した信号について有音部と無音部との識別を行なう有音/無音識別回路と、前記復号した情報により無音部を分類する雑音分類回路と、前記有音/無音識別回路で識別された前記無音部にて前記情報のうち前記ゲインと前記フィルタ係数の少なくとも一つに対して前記雑音分類回路による前記分類に応じて平滑化処理を施す平滑化回路と、前記平滑化回路によって平滑化処理を施した結果を用いて前記音声信号の復号を行う復号回路を含んで構成されることを特徴とする音声信号復号装置。

【請求項 9】 前記フィルタ係数とその長時間平均との差分に基づく変動量を長時間平均した値を用いて前記識別と前記分類の少なくとも一つを行なうことを特徴とする請求項 7 または 8 記載の音声信号復号装置。

【請求項 10】 前記受信した信号から入力信号のピッチ周期性とパワーの情報を復号し、前記識別及び前記分類の少なくとも一つにて前記ピッチ周期性

と前記パワーの情報の少なくとも一つを用いることを特徴とする請求項 7, 8 または 9 記載の音声信号復号装置。

【請求項 1 1】 復号した前記励振信号と前記音声信号から入力信号のピッチ周期性とパワーの情報を推定し、前記識別及び前記分類の少なくとも一つにて前記推定したピッチ周期性と前記推定したパワーの情報の少なくとも一つを用いることを特徴とする請求項 7, 8 または 9 記載の音声信号復号方法。

【請求項 1 2】 前記復号した情報のうち前記フィルタ係数の情報から得られる値を予め定められた閾値と比較することにより前記無音部を分類することを特徴とする請求項 8 記載の音声信号復号装置。

【請求項 1 3】 入力音声を少なくとも音源信号とゲインとフィルタ係数とで表現することにより符号化を行い、請求項 1、2、3、4、5 または 6 記載の音声信号復号方法で復号を行う音声信号符号化復号方法。

【請求項 1 4】 入力音声を少なくとも音源信号とゲインとフィルタ係数とで表現することにより符号化を行う音声信号符号化装置と、請求項 7、8、9、10、11 または 12 記載の音声信号復号装置を含んで構成されることを特徴とする音声信号符号化復号装置。

#### 【発明の詳細な説明】

##### 【0001】

#### 【発明の属する技術分野】

本発明は、音声信号を低ビットレートとするための符号化装置および復号装置に関し、特に、無音部での音質を改善する音声信号復号方法、音声信号復号装置、音声信号符号化復号方法及び音声信号符号化復号装置に関する。

##### 【0002】

#### 【従来技術】

音声信号を中低ビットレートで高能率に符号化する方法として、音声信号を線形予測フィルタとその駆動音源信号（音源信号）に分離して符号化する方法が広く用いられている。その代表的な方法の一つに CELP（Code Excited Linear Prediction）がある。CELP では、入力音声の周波数特性を表す線形予測係数が設定された線形予測フィルタを、音声のピッチ周期を表すピッチ信号と乱数やパルスか

ら成る音源信号との和で表される励振信号により駆動することで、合成音声信号（再生信号）が得られる。CELPに関してはM. Schroederらによる「Code excited linear prediction: High quality speech at very low bit rates」 (Proc. of IEEE Int. Conf. on Acoust., Speech and Signal Processing, pp.937-940, 1985)（文献1）を参照できる。

【0003】

携帯電話などの移動体通信では、繁華街の雑踏や走行中の自動車内に代表される雑音環境下での良好な通話品質が要求されるが、前述のCELPに基づく音声符号化では、雑音が重畳した音声（背景雑音音声）に対する音質が著しく劣化することが問題となっている。背景雑音音声の符号化品質改善に関して、復号器において音源信号のゲインを平滑化する方法が提案されている。

【0004】

音源信号のゲインを平滑化する方法に関しては、「Digital Cellular Telecommunication System; Adaptive Multi-Rate Speech Transcoding」 (ETSI Technical Report, GSM 06.90 version 2.0.0)（文献2）を参照できる。



図4は、音源信号のゲインを平滑化することで背景雑音音声の符号化品質を改善する、従来の音声信号復号装置の一例を示すブロック図である。ビット系列の入力は、 $T_{fr}$  msec (例えば、20 msec) 周期 (フレーム) で行われるものとし、再生ベクトルの計算は、 $N_{sfr}$  を整数 (例えば、4) として、 $T_{fr}/N_{sfr}$  msec (例えば、5 msec) 周期 (サブフレーム) で行われるものとする。フレーム長を  $L_{fr}$  サンプル (例えば、320 サンプル)、サブフレーム長を  $L_{sfr}$  サンプル (例えば、80 サンプル) とする。これらのサンプル数は、入力信号のサンプリング周波数 (例えば、16 kHz) によって定まる。以下、各ブロックを説明する。

入力端子10からビット系列の符号を入力する。符号入力回路1010は、入力端子10から入力したビット系列の符号を分割し、複数の復号パラメータに対応するインデックスに変換する。そして、入力信号の周波数特性を表す線スペクトル対 (Line Spectrum Pair, LSP) に対応するインデックスを LSP 復号回路1020へ出力し、入力信号のピッチ周期を表す遅延  $L_{pd}$  に対応するインデックスをピッチ信号復号回路1210へ出力し、乱数やパルスから成る音源ベクトルに対応するインデックスを音源信号復号回路1110に出力し、第1のゲインに対応するインデックスを第1のゲイン復号回路1220に出力し、第2のゲインに対応するインデックスを第2のゲイン復号回路1120に出力する。LSP 復号回路1020は、複数セットの LSP が格納されたテーブルを備えている。LSP 復号回路1020は、符号入力回路1010から出力されるインデックスを入力し、前記インデックスに対応する LSP を前記テーブルより読み出し、現フレーム (第  $n$  フレーム) の第  $N_{sfr}$  サブフレームにおける  $LSP \hat{q}_j^{(N_{\phi})}(n)$ ,  $j=1, \Lambda, N_p$  とする。ここで、 $N_p$  は線形予測次数である。第1から第  $N_{sfr}-1$  サブフレームの LSP は、 $\hat{q}_j^{(N_{\phi})}(n)$  と  $\hat{q}_j^{(N_{\phi})}(n-1)$  とを線形補間して求める。前記  $LSP \hat{q}_j^{(m)}(n)$ ,  $j=1, \Lambda, N_p$ ,  $m=1, \Lambda, N_{sfr}$  を線形予測係数変換回路1030および平滑化係数計算回路1310へ出力する。線形予測係数変換回路1030は、LSP 復号回路1020から出力された  $LSP \hat{q}_j^{(m)}(n)$ ,  $j=1, \Lambda, N_p$ ,  $m=1, \Lambda, N_{sfr}$  を入力する。入力された前記  $\hat{q}_j^{(m)}(n)$  を線形予測係数  $\hat{\alpha}_j^{(m)}(n)$ ,  $j=1, \Lambda, N_p$ ,  $m=1, \Lambda, N_{sfr}$  に変換し、前記  $\hat{\alpha}_j^{(m)}(n)$  を合成フィルタ1040へ出力する。ここで、LSP から線形予測係数への変換に関しては周知の方法、例えば、文献2の5.2.4節に記述されている方法を参照できる。

【0005】

音源信号復号回路 1 1 1 0 は、複数個の音源ベクトルが格納されたテーブルを備えている。音源信号復号回路 1 1 1 0 は、符号入力回路 1 0 1 0 から出力されるインデックスを入力し、前記インデックスに対応する音源ベクトルを、前記テーブルより読み出し、第 2 のゲイン回路 1 1 2 0 へ出力する。

第 2 のゲイン復号回路 1 1 2 0 は、複数個のゲインが格納されたテーブルを備えている。第 2 のゲイン復号回路 1 1 2 0 は、符号入力回路 1 0 1 0 から出力されるインデックスを入力し、前記インデックスに対応する第 2 のゲインを前記テーブルより読み出し、平滑化回路 1 3 2 0 へ出力する。第 2 のゲイン回路 1 1 3 0 は、音源信号復号回路 1 1 1 0 から出力される第 1 の音源ベクトルと、平滑化回路 1 3 2 0 から出力される第 2 のゲインとを入力し、前記第 1 の音源ベクトルと前記第 2 のゲインとを乗算し、第 2 の音源ベクトルを復号し、復号した前記第 2 の音源ベクトルを加算器 1 0 5 0 へ出力する。記憶回路 1 2 4 0 は、加算器 1 0 5 0 から励振ベクトルを入力し、保持する。記憶回路 1 2 4 0 は、過去に入力されて保持されている前記励振ベクトルをピッチ信号復号回路 1 2 1 0 へ出力する。ピッチ信号復号回路 1 2 1 0 は、記憶回路 1 2 4 0 に保持されている過去の励振ベクトルと符号入力回路 1 0 1 0 から出力されるインデックスとを入力する。前記インデックスは、遅延  $L_{pd}$  を指定する。そして、前記過去の励振ベクトルにおいて、現フレームの始点より  $L_{pd}$  サンプル過去の点から、ベクトル長に相当する  $L_{sf}$  サンプル分のベクトルを切り出し、第 1 のピッチ信号(ベクトル)を復号する。ここで、 $L_{pd} < L_{sf}$  の場合には  $L_{pd}$  サンプル分のベクトルを切り出し、この切り出した  $L_{pd}$  サンプルを繰り返し接続して、ベクトル長が  $L_{sf}$  サンプルである第 1 のピッチベクトルを復号する。

【0006】

ピッチ信号復号回路1210は、前記第1のピッチベクトルを第1のゲイン回路1230へ出力する。第1のゲイン復号回路1220は、複数のゲインが格納されたテーブルを備えている。第1のゲイン復号回路1220は、符号入力回路1010から出力されるインデックスを入力し、前記インデックスに対応する第1のゲインを前記テーブルより読み出し、第1のゲイン回路1230へ出力する。第1のゲイン回路1230は、ピッチ信号復号回路1210から出力される第1のピッチベクトルと、第1のゲイン復号回路1220から出力される第1のゲインとを入力し、前記第1のピッチベクトルと前記第1のゲインとを乗算し、第2のピッチベクトルを生成し、生成した前記第2のピッチベクトルを加算器1050へ出力する。加算器1050は、第1のゲイン回路1230から出力される第2のピッチベクトルと、第2のゲイン回路1130から出力される第2の音源ベクトルを入力し、これらの和を計算し、これを励振ベクトルとして、合成フィルタ1040へ出力する。平滑化係数計算回路1310は、LSP復号回路1020から出力されるLSP  $\hat{q}_j^{(m)}(n)$  を入力し、平均LSP  $\bar{q}_{0j}(n)$  を次式により計算する。

$$\bar{q}_{0j}(n) = 0.84 \cdot \bar{q}_{0j}(n-1) + 0.16 \cdot \hat{q}_j^{(N_{\phi})}(n)$$

次に、各サブフレーム  $m$  に対して、LSP の変動量  $d_0(m)$  を次式により計算する。

$$d_0(m) = \sum_{j=1}^{N_2} \frac{|\bar{q}_{0j}(n) - \hat{q}_j^{(m)}(n)|}{\bar{q}_{0j}(n)}$$

サブフレーム  $m$  の平滑化係数  $k_0(m)$  は、次式で計算される。

$$k_0(m) = \min(0.25, \max(0, d_0(m) - 0.4)) / 0.25$$

ここで、 $\min(x, y)$  は  $x$  と  $y$  のうち小さい方を、 $\max(x, y)$  は  $x$  と  $y$  のうち大きい方を値としてとる関数である。最後に、前記平滑化係数  $k_0(m)$  を平滑化回路1320へ出力する。平滑化回路1320は、平滑化係数計算回路1310から出力される平滑化係数  $k_0(m)$  と、第2のゲイン復号回路1120から出力される第2のゲインとを入力する。サブフレーム  $m$  の第2のゲイン  $\hat{g}_0(m)$  から平均ゲイン  $\bar{g}_0(m)$  を次式により計算する。

【0 0 0 7】

$$\bar{g}_0(m) = \frac{1}{5} \sum_{i=0}^4 \hat{g}_0(m-i)$$

次に、次式により第2のゲイン $\hat{g}_0(m)$ は置き換えられる。

$$\hat{g}_0(m) = \hat{g}_0(m) \cdot k_0(m) + \bar{g}_0(m) \cdot (1 - k_0(m))$$

最後に、前記第2のゲイン $\hat{g}_0(m)$ を第2のゲイン回路1 1 3 0に出力する。

合成フィルタ1 0 4 0は、加算器1 0 5 0から出力される励振ベクトルと、線形予測係数変換回路1 0 3 0から出力される線形予測係数 $\alpha_i, i=1, \Lambda, N_p$ を入力する。線形予測係数が設定された合成フィルタ $1/A(z)$ を、前記励振ベクトルにより駆動することで、再生ベクトルを計算し、出力端子2 0から出力する。ここで、合成フィルタの伝達関数 $1/A(z)$ は、線形予測係数を $\alpha_i, i=1, \Lambda, N_p$ と

すると、 $1/A(z) = 1/(1 - \sum_{i=1}^{N_p} \alpha_i' z^i)$ と表される。

図5は、従来例による音声信号符号化復号装置における、音声信号符号化装置の構成を示すブロック図である。なお、第1のゲイン回路1 2 3 0、第2のゲイン回路1 1 3 0、加算器1 0 5 0および記憶回路1 2 4 0は、図4の従来例における音声信号復号装置で説明した各ブロックと同じであるので、ここでは説明を省略する。音声信号をサンプリングし、この複数サンプルを1フレームとして一つのベクトルにまとめて生成した入力信号（入力ベクトル）を入力端子3 0から入力する。線形予測係数計算回路5 5 1 0は、入力端子3 0から入力ベクトルを入力する。前記入力ベクトルに対して線形予測分析を行い線形予測係数を求める。線形予測分析に関しては、周知の方法、例えば、文献4の第8章「Linear Predictive Coding of Speech」を参照できる。線形予測係数計算回路5 5 1 0は、前記線形予測係数を、LSP変換/量子化回路5 5 2 0と、重み付けフィルタ5 0 5 0と、重み付け合成フィルタ5 0 4 0とへ出力する。LSP変換/量子化回路5 5 2 0は、線形予測係数計算回路5 5 1 0から出力される線形予測係数を入力し、前記線形予測係数をLSPへ変換し、前記LSPを量子化して量子化LSPを得る。ここで、線形予測係数からLSPへの変換に関しては、周知の方法、例えば、文献2の5.2.4節に記述されている方法を参照できる。

【0008】

また、LSP の量子化に関しては、文献 2 の 5.2.5 節に記述されている方法を参照できる。また、前記量子化 LSP は、図 4 (従来例) の LSP 復号回路で説明したように、現フレーム (第  $n$  フレーム) の第  $N_{\text{eff}}$  サブフレームにおける量子化 LSP  $\hat{q}_j^{(N_{\text{eff}})}(n)$ ,  $j=1, \Lambda, N_p$  とする。そして、第 1 から第  $N_{\text{eff}}-1$  サブフレームの量子化 LSP は、 $\hat{q}_j^{(N_{\text{eff}})}(n)$  と  $\hat{q}_j^{(N_{\text{eff}})}(n-1)$  とを線形補間して求める。さらに、前記 LSP は、現フレーム (第  $n$  フレーム) の第  $N_{\text{eff}}$  サブフレームにおける LSP  $\hat{q}_j^{(N_{\text{eff}})}(n)$ ,  $j=1, \Lambda, N_p$  とする。そして、第 1 から第  $N_{\text{eff}}-1$  サブフレームの LSP は、 $q_j^{(N_{\text{eff}})}(n)$  と  $q_j^{(N_{\text{eff}})}(n-1)$  とを線形補間して求める。

LSP 変換 / 量子化回路 5 5 2 0 は、前記 LSP  $q_j^{(m)}(n)$ ,  $j=1, \Lambda, N_p$ ,  $m=1, \Lambda, N_{\text{eff}}$  と、前記量子化 LSP  $\hat{q}_j^{(m)}(n)$ ,  $j=1, \Lambda, N_p$ ,  $m=1, \Lambda, N_{\text{eff}}$  と、を線形予測係数変換回路 5 0 3 0 へ出力し、前記量子化 LSP  $\hat{q}_j^{(N_{\text{eff}})}(n)$ ,  $j=1, \Lambda, N_p$  に対応するインデックスを符号出力回路 6 0 1 0 へ出力する。線形予測係数変換回路 5 0 3 0 は、LSP 変換/量子化回路 5 5 2 0 から出力される LSP  $q_j^{(m)}(n)$ ,  $j=1, \Lambda, N_p$ ,  $m=1, \Lambda, N_{\text{eff}}$  と量子化 LSP  $\hat{q}_j^{(m)}(n)$ ,  $j=1, \Lambda, N_p$ ,  $m=1, \Lambda, N_{\text{eff}}$  とを入力し、前記  $q_j^{(m)}(n)$  を線形予測係数  $\alpha_j^{(m)}(n)$ ,  $j=1, \Lambda, N_p$ ,  $m=1, \Lambda, N_{\text{eff}}$  に変換し、前記  $\hat{q}_j^{(m)}(n)$  を量子化線形予測係数  $\hat{\alpha}_j^{(m)}(n)$ ,  $j=1, \Lambda, N_p$ ,  $m=1, \Lambda, N_{\text{eff}}$  に変換し、前記  $\alpha_j^{(m)}(n)$  を重み付けフィルタ 5 0 5 0 と重み付け合成フィルタ 5 0 4 0 とへ出力し、前記  $\hat{\alpha}_j^{(m)}(n)$  を重み付け合成フィルタ 5 0 4 0 へ出力する。ここで、LSP から線形予測係数への変換および量子化 LSP から量子化線形予測係数への変換に関しては周知の方法、例えば、文献 2 の 5.2.4 節に記述されている方法を参照できる。重み付けフィルタ 5 0 5 0 は、入力端子 3 0 から入力ベクトルを入力し、線形予測係数変換回路 5 0 3 0 から出力される線形予測係数を入力し、前記線形予測係数を用いて、人間の聴覚特性に対応した重みづけフィルタ  $W(z)$  を生成し、前記重みづけフィルタを前記入力ベクトルで駆動することで、重みづけ入力ベクトルを得る。そして前記重みづけ入力ベクトルを、差分器 5 0 6 0 へと出力する。ここで、重みづけフィルタの伝達関数  $W(z)$  は、 $W(z)=Q(z/\gamma_1)/Q(z/\gamma_2)$  と表される。

【0009】

ただし、 $Q(z/\gamma_1)=1-\sum_{i=1}^{N_p}\alpha_i^{(m)}\gamma_1^iz^i$ 、 $Q(z/\gamma_2)=1-\sum_{i=1}^{N_p}\alpha_i^{(m)}\gamma_2^iz^i$ である。 $\gamma_1$ および $\gamma_2$ は定数であり、例えば、 $\gamma_1=0.9$ 、 $\gamma_2=0.6$ である。また、重みづけフィルタの詳細に関しては、(文献1)を参照できる。重み付け合成フィルタ5040は、加算器1050から出力される励振ベクトルと、線形予測係数変換回路5030から出力される線形予測係数 $\alpha_j^{(m)}(n)$ 、 $j=1, \Lambda, N_p$ 、 $m=1, \Lambda, N_{sp}$ と、量子化線形予測係数 $\hat{\alpha}_j^{(m)}(n)$ 、 $j=1, \Lambda, N_p$ 、 $m=1, \Lambda, N_{sp}$ と、を入力する。 $\alpha_j^{(m)}(n)$ と $\hat{\alpha}_j^{(m)}(n)$ が設定された重み付け合成フィルタ $H(z)W(z)=Q(z/\gamma_1)/[A(z)Q(z/\gamma_2)]$ を、前記励振ベクトルにより駆動することで、重み付け再生ベクトルを得る。ここで、合成フィルタの伝達関数 $H(z)=1/A(z)$ は、 $1/A(z)=1/(1-\sum_{i=1}^{N_p}\hat{\alpha}_i^{(m)}z^i)$ と表される。差分器5060は、重み付けフィルタ5050から出力される重み付け入力ベクトルを入力し、重み付け合成フィルタ5040から出力される重み付け再生ベクトルを入力し、それらの差分を計算し、これを差分ベクトルとして、最小化回路5070へ出力する。

最小化回路5070は、音源信号生成回路5110に格納されている音源ベクトル全てに対応するインデックスを、前記音源生成回路5110へ順次出力し、ピッチ信号生成回路5210において規定された範囲内の遅延 $L_{pd}$ 全てに対応するインデックスを、前記ピッチ信号生成回路5210へ順次出力し、第1のゲイン生成回路6220に格納されている第1のゲイン全てに対応するインデックスを、前記第1のゲイン生成回路6220へ順次出力し、第2のゲイン生成回路6120に格納されている第2のゲイン全てに対応するインデックスを、前記第2のゲイン生成回路6120へ順次出力する。また、差分器5060から出力される差分ベクトルを順次入力し、そのノルムを計算し、前記ノルムが最小となるような、前記音源ベクトル、前記遅延 $L_{pd}$ 、前記第1のゲインおよび前記第2のゲインを選択し、これらに対応するインデックスを符号出力回路6010へ出力する。ピッチ信号生成回路5210、音源信号生成回路5110、第1のゲイン生成回路6220および第2のゲイン生成回路6120は、各々、最小化回路5070から出力されるインデックスを順次入力する。

## 【0 0 1 0】

これら、ピッチ信号生成回路 5 2 1 0、音源信号生成回路 5 1 1 0、第 1 のゲイン生成回路 6 2 2 0 および第 2 のゲイン生成回路 6 1 2 0 は、各々、入出力に関する結線を除けば、図 4 のピッチ信号復号回路 1 2 1 0、音源信号復号回路 1 1 1 0、第 1 のゲイン復号回路 1 2 2 0 および第 2 のゲイン復号回路 1 1 2 0 と同じであるので、これら各ブロックの詳細な説明は省略する。符号出力回路 6 0 1 0 は、LSP 変換/量子化回路 5 5 2 0 から出力される量子化 LSP に対応するインデックスを入力し、最小化回路 5 0 7 0 から出力される、音源ベクトル、遅延  $L_{pd}$ 、第 1 のゲインおよび第 2 のゲインの各々に対応するインデックスを入力し、各インデックスをビット系列の符号に変換し、出力端子 4 0 を介して出力する。

## 【0 0 1 1】

## 【発明が解決しようとする課題】

第 1 の問題点は、有音部で間欠的に含まれる短い無音部や有音部の一部において通常の有音部とは異質の音を生成する場合があります、その結果、有音部において不連続音が発生することである。その理由は、この短い無音部において、LSP の変動量  $d_0(m)$  が小さくなるため、平滑化係数が大きくなるためである。また、 $d_0(m)$  の時間変動(ばらつき)が大きいため、有音部の一部において、 $d_0(m)$  がある程度大きい値を持つが、平滑化係数が零とはならないためである。

第 2 の問題点は、無音部において平滑化係数が急激に変化する場合があります、その結果、無音部において不連続音が発生する。その理由は、時間的な変動の大きい  $d_0(m)$  を用いて平滑化係数を定めているためである。

第 3 の問題点は、背景雑音の種類に応じた適切な平滑化処理を選択できない場合があります、その結果、復号音質が劣化することである。その理由は、設定パラメータが異なるだけの単一アルゴリズムに基づいて復号パラメータを平滑化するからである。

## 【0 0 1 2】

## 【課題を解決するための手段】

本願の第 1 の発明は、受信した信号から少なくとも音源信号とゲインとフィルタ係数の情報を復号し、前記音源信号に前記ゲインを乗じて得られる励振信号により前記フィルタ係数で構成するフィルタを駆動することによって音声信号を復号する音声信号復号方法において、前記復号した情報を用いて前記受信した信号について有音部と無音部との識別を行ない、前記無音部にて前記情報のうち前記ゲインと前記フィルタ係数の少なくとも一つに対して前記情報に基づく平滑化処理を施した結果を用いて前記音声信号の復号を行うことを特徴とする。

## 【 0 0 1 3 】

本願の第 2 の発明は、受信した信号から少なくとも音源信号とゲインとフィルタ係数の情報を復号し、前記音源信号に前記ゲインを乗じて得られる励振信号により前記フィルタ係数で構成するフィルタを駆動することによって音声信号を復号する音声信号復号方法において、前記復号した情報を用いて前記受信した信号について有音部と無音部との識別を行うとともに前記復号した情報により無音部を分類し、前記無音部にて前記情報のうち前記ゲインと前記フィルタ係数の少なくとも一つに対して前記分類に応じて平滑化処理を施した結果を用いて前記音声信号の復号を行うことを特徴とする。

## 【 0 0 1 4 】

本願の第 3 の発明は、第 1 の発明または第 2 の発明において、前記フィルタ係数とその長時間平均との差分に基づく変動量を長時間平均した値を用いて前記識別と前記分類の少なくとも一つを行なうことを特徴とする。

本願の第 4 の発明は、第 1、第 2 の発明または第 3 の発明において、前記受信した信号から入力信号のピッチ周期性和パワーの情報を復号し、前記識別及び前記分類の少なくとも一つにて前記ピッチ周期性和前記パワーの情報の少なくとも一つを用いることを特徴とする。

## 【 0 0 1 5 】

本願の第 5 の発明は、第 1、第 2 の発明または第 3 の発明において、復号した前記励振信号と前記音声信号から入力信号のピッチ周期性和パワーの情報を推定し、前記識別及び前記分類の少なくとも一つにて前記推定したピッチ周期性和前記推定したパワーの情報の少なくとも一つを用いることを特徴とする。



## 【0016】

本願の第6の発明は、第2の発明において、前記復号した情報のうち前記フィルタ係数の情報から得られる値を予め定められた閾値と比較することにより前記無音部を分類することを特徴とする。

## 【0017】

本願の第7の発明は、受信した信号から少なくとも音源信号とゲインとフィルタ係数の情報を復号し、前記音源信号に前記ゲインを乗じて得られる励振信号により前記フィルタ係数で構成するフィルタを駆動することによって音声信号を復号する音声信号復号装置において、前記復号した情報を用いて前記受信した信号について有音部と無音部との識別を行なう有音/無音識別回路と、前記有音/無音識別回路で識別された前記無音部にて前記情報のうち前記ゲインと前記フィルタ係数の少なくとも一つに対して前記情報に基づく平滑化処理を施す平滑化回路と、前記平滑化回路によって平滑化処理を施した結果を用いて前記音声信号の復号を行う復号回路を含んで構成されることを特徴とする。

## 【0018】

本願の第8の発明は、受信した信号から少なくとも音源信号とゲインとフィルタ係数の情報を復号し、前記音源信号に前記ゲインを乗じて得られる励振信号により前記フィルタ係数で構成するフィルタを駆動することによって音声信号を復号する音声信号復号装置において、前記復号した情報を用いて前記受信した信号について有音部と無音部との識別を行なう有音/無音識別回路と、前記復号した情報により無音部を分類する雑音分類回路と、前記有音/無音識別回路で識別された前記無音部にて前記情報のうち前記ゲインと前記フィルタ係数の少なくとも一つに対して前記雑音分類回路による前記分類に応じて平滑化処理を施す平滑化回路と、前記平滑化回路によって平滑化処理を施した結果を用いて前記音声信号の復号を行う復号回路を含んで構成されることを特徴とする。

## 【0019】

本願の第9の発明は、第7の発明または第8の発明において、前記フィルタ係数とその長時間平均との差分に基づく変動量を長時間平均した値を用いて前記識別と前記分類の少なくとも一つを行なうことを特徴とする。

本願の第 1 0 の発明は、第 7、第 8 の発明または第 9 の発明において、前記受信した信号から入力信号のピッチ周期性とパワーの情報を復号し、前記識別及び前記分類の少なくとも一つにて前記ピッチ周期性と前記パワーの情報の少なくとも一つを用いることを特徴とする。

【0 0 2 0】

本願の第 1 1 の発明は、第 7、第 8 の発明または第 9 の発明において、復号した前記励振信号と前記音声信号から入力信号のピッチ周期性とパワーの情報を推定し、前記識別及び前記分類の少なくとも一つにて前記推定したピッチ周期性と前記推定したパワーの情報の少なくとも一つを用いることを特徴とする。

本願の第 1 2 の発明は、第 8 の発明において、前記復号した情報のうち前記フィルタ係数の情報から得られる値を予め定められた閾値と比較することにより前記無音部を分類することを特徴とする。

【0 0 2 1】

本願の第 1 3 の発明は、入力音声を少なくとも音源信号とゲインとフィルタ係数とで表現することにより符号化を行い、請求項 1、2、3、4、5 または 6 記載の音声信号復号方法で復号を行うことを特徴とする。

本願の第 1 4 の発明は、入力音声を少なくとも音源信号とゲインとフィルタ係数とで表現することにより符号化を行う音声信号符号化装置と、請求項 7、8、9、1 0、1 1 または 1 2 記載の音声信号復号装置を含んで構成されることを特徴とする。

【0 0 2 2】

〔作用〕

本発明において、 $d_0(m)$ の長時間平均値は、 $d_0(m)$ に比べて時間的な変動が緩やかであるため、有音部において間欠的に小さくならない。従って、この平均値に応じて平滑化係数を決定すれば、有音部に間欠的に含まれる短い無音部により生じていた不連続音を削減できる。また、この平均値等を用いて有音部か無音部であるかの識別を行うことにより、有音部では完全に復号パラメータの平滑化係数を零とすることが可能になる。

次に、無音部に関しても同様に、 $d_0(m)$ の長時間平均値を用いることにより、平滑化係数が急激に変化することを回避することが可能である。

最後に、本発明では、復号パラメータを単一の処理ではなく、無音部において、入力信号の特性を考慮して複数個用意した処理方法を選択的に用いて復号パラメータの平滑化を行う。例えば、過去の限られた区間の復号パラメータから計算する移動平均(Moving Average)処理、長い過去の影響を考慮可能な自己回帰(Auto-Regressive)処理、あるいは、平均を計算した後に予め設定した値を上限や下限で制限する非線形処理等がある。

【 0 0 2 3 】

## 【発明の実施の形態】

次に、本発明の実施の形態について図面を参照して詳細に説明する。

図1は、本発明の音声信号復号装置の第1の実施の形態を示すブロック図である。

なお、入力端子10、出力端子20、LSP復号回路1020、線形予測係数変換回路1030、音源信号復号回路1110、記憶回路1240、ピッチ信号復号回路1210、第1のゲイン回路1230、第2のゲイン回路1130、加算器1050および合成フィルタ1040は、図4の従来例で説明した各ブロックと同じであるので、ここでは説明を省略する。以下では、符号入力回路1010、有音/無音識別回路2020、雑音分類回路2030、第1の切替回路2110、第2の切替回路2210、第1のフィルタ2150、第2のフィルタ2160、第3のフィルタ2170、第4のフィルタ2250、第5のフィルタ2260、第6のフィルタ2270、第1のゲイン復号回路2220および第2のゲイン復号回路2120について説明する。ここで、ビット系列の入力は、 $T_{fr}$  msec (例えば、20 msec) 周期 (フレーム) で行われるものとし、再生ベクトルの計算は、 $N_{fr}$  を整数 (例えば、4) として、 $T_{fr}/N_{fr}$  msec (例えば、5 msec) 周期 (サブフレーム) で行われるものとする。フレーム長を  $L_{fr}$  サンプル (例えば、320 サンプル)、サブフレーム長を  $L_{sf}$  サンプル (例えば、80 サンプル) とする。これは、入力信号のサンプリング周波数 (例えば、16 kHz) によって定まる。以下、各ブロックを説明する。

符号入力回路1010は、入力端子10から入力したビット系列の符号を分割し、複数の復号パラメータに対応するインデックスに変換する。そして、LSPに対応するインデックスをLSP復号回路1020へ出力し、音声モードに対応するインデックスを音声モード復号回路2050へ出力し、フレームエネルギーに対応するインデックスをフレームパワー復号回路2040へ出力し、遅延  $L_{pd}$  に対応するインデックスをピッチ信号復号回路1210へ出力し、音源ベクトルに対応するインデックスを音源信号復号回路1110へ出力し、第1のゲインに対応するインデックスを第1のゲイン復号回路2220へ出力し、第2のゲインに対応するインデックスを第2のゲイン復号回路2120へ出力する。

【 0 0 2 4 】

音声モード復号回路 2 0 5 0 は、符号入力回路 1 0 1 0 から出力される音声モードに対応するインデックスを入力し、前記インデックスに対応する音声モード  $S_{mode}$  を設定する。音声モードは、音声符号化器において、聴感重み付け後の入力信号を用いて求められる開ループピッチ予測ゲイン  $G_{op}(m)$  のフレーム内平均値  $\bar{G}_{op}(n)$  を閾値処理することで決定され、復号器に伝送される。ここで、 $n$  はフレーム番号、 $m$  はサブフレーム番号を表す。また、音声モードの決定に関しては、K. Ozawa らによる「M-LCELP Speech Coding at 4 kb/s with Multi-Mode and Multi-Codebook」(IEICE Trans. On Commun., Vol. E77-B, No. 9, pp.1114-1121, 1994) (文献 3) を参照できる。音声モード復号回路 2 0 5 0 は、前記音声モード  $S_{mode}$  を有音/無音識別回路 2 0 2 0 と第 1 のゲイン復号回路 2 2 2 0 と第 2 のゲイン復号回路 2 1 2 0 とへ出力する。フレームパワー復号回路 2 0 4 0 は、複数のフレームエネルギーが格納されたテーブルを備えている。符号入力回路 1 0 1 0 から出力されるフレームパワーに対応するインデックスを入力し、前記インデックスに対応するフレームパワー  $\hat{E}_{rms}$  を前記テーブルより読み出す。ここで、フレームパワーは、音声符号化器において、入力信号のパワーを量子化して求められ、その量子化値に対応するインデックスが復号器へ伝送される。フレームパワー復号回路 2 0 4 0 は、前記フレームパワー  $\hat{E}_{rms}$  を有音/無音識別回路 2 0 2 0 と第 1 のゲイン復号回路 2 2 2 0 と第 2 のゲイン復号回路 2 1 2 0 とへ出力する。有音/無音識別回路 2 0 2 0 は、LSP 復号回路 1 0 2 0 から出力される LSP  $\hat{q}_j^{(m)}(n)$  と、音声モード復号回路 2 0 5 0 から出力される音声モード  $S_{mode}$  と、フレームパワー復号回路 2 0 4 0 から出力されるフレームパワー  $\hat{E}_{rms}$  とを入力する。スペクトルパラメータの変動量を求める手順を以下に示す。スペクトルパラメータとして LSP  $\hat{q}_j^{(m)}(n)$  を用いる。第  $n$  フレームにおいて、LSP の長時間平均  $\bar{q}_j(n)$  を次式により計算する。

$$\bar{q}_j(n) = \beta_0 \cdot \bar{q}_j(n-1) + (1 - \beta_0) \cdot \hat{q}_j^{(N_p)}(n), \quad j = 1, \Lambda, N_p$$

ここで、 $\beta_0 = 0.9$  である。第  $n$  フレームにおける LSP の変動量  $d_q(n)$  を次式により定義する。

【 0 0 2 5 】

$$d_q(n) = \sum_{j=1}^{N_p} \sum_{m=1}^{N_{sp}} \frac{D_{q,j}^{(m)}(n)}{\bar{q}_j(n)}$$

ここで、 $D_{q,j}^{(m)}(n)$ は、 $\bar{q}_j(n)$ と $\hat{q}_j^{(m)}(n)$ との距離に相当する。例えば、

$$D_{q,j}^{(m)}(n) = (\bar{q}_j(n) - \hat{q}_j^{(m)}(n))^2$$

または、

$$D_{q,j}^{(m)}(n) = |\bar{q}_j(n) - \hat{q}_j^{(m)}(n)|$$

とできる。ここでは、 $D_{q,j}^{(m)}(n) = |\bar{q}_j(n) - \hat{q}_j^{(m)}(n)|$ とする。前記変動量 $d_q(n)$ が大きい区間を有音部に、小さい区間を無音部に概ね対応させることができる。しかし、前記変動量 $d_q(n)$ は、時間的な変動が大きく、有音部における $d_q(n)$ の値域と、無音部における $d_q(n)$ の値域は、互いに重複するため、有音部と無音部とを識別するための閾値の設定が容易ではない問題がある。そこで、 $d_q(n)$ の長時間平均を有音部と無音部との識別に用いる。線形フィルタまたは非線形フィルタを用いて $d_q(n)$ の長時間平均 $\bar{d}_{q1}(n)$ を求める。 $\bar{d}_{q1}(n)$ には、例えば、 $d_q(n)$ の平均値、中央値、最頻値などが適用できる。ここでは、次式を用いる。

$$\bar{d}_{q1}(n) = \beta_1 \cdot \bar{d}_{q1}(n-1) + (1 - \beta_1) \cdot d_q(n)$$

ここで、 $\beta_1 = 0.9$ である。

$\bar{d}_{q1}(n)$ に対する閾値処理により、識別フラグ $S_{vs}$ を決定する。

if  $(\bar{d}_{q1}(n) \geq C_{th1})$  then  $S_{vs} = 1$

else  $S_{vs} = 0$

ここで、 $C_{th1}$ はある定数（例えば、2.2）であり、 $S_{vs} = 1$ は有音部に、 $S_{vs} = 0$ は無音部に対応する。有音部でも定常性が高い区間では、 $d_q(n)$ が小さいため、無音部と誤る場合がある。そのため、フレームパワーが大きく、かつピッチ予測ゲインが大きい場合には有音部とみなすこととする。 $S_{vs} = 0$ のとき、次の追加判定により、 $S_{vs}$ の修正を行う。

if  $(\hat{E}_{ms} \geq C_{ms} \text{ and } S_{mode} \geq 2)$  then  $S_{vs} = 1$

else  $S_{vs} = 0$

ここで、 $C_{ms}$ はある定数（例えば、10000）である。 $S_{mode} \geq 2$ は、ピッチ予測ゲインのフレーム内平均値 $\bar{G}_{op}(n)$ が 3.5 dB 以上であることに対応する。

【 0 0 2 6 】

これは符号化器において定義される。有音/無音識別回路 2 0 2 0 は、 $S_{vs}$  を雑音分類回路 2 0 3 0 と、第 1 の切替回路 2 1 1 0 および第 2 の切替回路 2 2 1 0 とへ出力し、 $\bar{d}_{q1}(n)$  を雑音分類回路 2 0 3 0 へ出力する。雑音分類回路 2 0 3 0 は、有音/無音識別回路 2 0 2 0 から出力される  $\bar{d}_{q1}(n)$  と  $S_{vs}$  を入力する。無音部（雑音部）において、線形フィルタまたは非線型フィルタを用いて  $\bar{d}_{q1}(n)$  の平均的な挙動を反映した値  $\bar{d}_{q2}(n)$  を求める。 $S_{vs}=0$  のとき次式を計算する。

$$\bar{d}_{q2}(n) = \beta_2 \cdot \bar{d}_{q2}(n-1) + (1 - \beta_2) \cdot \bar{d}_{q1}(n)$$

ここで、 $\beta_2 = 0.94$  である。 $\bar{d}_{q2}(n)$  に対する閾値処理により、雑音の分類を行い、分類フラグ  $S_{nz}$  を決定する。

$$\text{if } (\bar{d}_{q2}(n) \geq C_{th2}) \text{ then } S_{nz} = 1$$

$$\text{else } S_{nz} = 0$$

ここで、 $C_{th2}$  はある定数（例えば、1.7）であり、 $S_{nz}=1$  は周波数特性の時間変化が非定常的である雑音に、 $S_{nz}=0$  は周波数特性の時間変化が定常的である雑音に対応する。雑音分類回路 2 0 3 0 は、 $S_{nz}$  を第 1 の切替回路 2 1 1 0 と第 2 の切替回路 2 2 1 0 へ出力する。

第 1 の切替回路 2 1 1 0 は、LSP 復号回路 1 0 2 0 から出力される LSP  $\hat{q}_j^{(m)}(n)$  と、有音/無音識別回路 2 0 2 0 から出力される識別フラグ  $S_{vs}$  と、雑音分類回路 2 0 3 0 から出力される分類フラグ  $S_{nz}$  とを入力し、前記識別フラグの値と分類フラグの値に応じてスイッチを切り替えることで、前記 LSP  $\hat{q}_j^{(m)}(n)$  を、 $S_{vs}=0$  かつ  $S_{nz}=0$  のときは第 1 のフィルタ 2 1 5 0 へ、 $S_{vs}=0$  かつ  $S_{nz}=1$  のときは第 2 のフィルタ 2 1 6 0 へ、 $S_{vs}=1$  のときは第 3 のフィルタ 2 1 7 0 へ出力する。第 1 のフィルタ 2 1 5 0 は、第 1 の切替回路 2 1 1 0 から出力される LSP  $\hat{q}_j^{(m)}(n)$  を入力し、線形フィルタまたは非線型フィルタを用いて平滑化し、これを第 1 の平滑化 LSP  $\bar{q}_{1,j}^{(m)}(n)$  とし、線形予測係数変換回路 1 0 3 0 へ出力する。ここでは、次式で表されるフィルタを用いる。

$$\bar{q}_{1,j}^{(m)}(n) = \gamma_1 \cdot \bar{q}_{1,j}^{(m-1)}(n) + (1 - \gamma_1) \cdot \hat{q}_j^{(m)}(n), \quad j = 1, \Lambda, N_p$$

ただし、 $\bar{q}_{1,j}^{(0)}(n) = \bar{q}_{1,j}^{(N_p)}(n-1)$  とする。

【0 0 2 7】

また、 $\gamma_1 = 0.5$ とする。第2のフィルタ 2 1 6 0は、第1の切替回路 2 1 1 0から出力されるLSP  $\hat{q}_j^{(m)}(n)$ を入力し、線形フィルタまたは非線型フィルタを用いて平滑化し、これを第2の平滑化LSP  $\bar{q}_{2,j}^{(m)}(n)$ とし、線形予測係数変換回路 1 0 3 0へ出力する。ここでは、次式で表されるフィルタを用いる。

$$\bar{q}_{2,j}^{(m)}(n) = \gamma_2 \cdot \bar{q}_{2,j}^{(m-1)}(n) + (1 - \gamma_2) \cdot \hat{q}_j^{(m)}(n), \quad j = 1, \Lambda, N_p$$

ただし、 $\bar{q}_{2,j}^{(0)}(n) = \bar{q}_{2,j}^{(N_p)}(n-1)$ とする。また、 $\gamma_2 = 0.0$ とする。第3のフィルタ 2 1 7 0は、第1の切替回路 2 1 1 0から出力されるLSP  $\hat{q}_j^{(m)}(n)$ を入力し、線形フィルタまたは非線型フィルタを用いて平滑化し、これを第3の平滑化LSP  $\bar{q}_{3,j}^{(m)}(n)$ とし、線形予測係数変換回路 1 0 3 0へ出力する。ここでは、 $\bar{q}_{3,j}^{(m)}(n) = \hat{q}_j^{(m)}(n)$ とする。第2の切替回路 2 2 1 0は、第2のゲイン復号回路 2 1 2 0から出力される第2のゲイン  $\hat{g}_2^{(m)}(n)$ と、有音/無音識別回路 2 0 2 0から出力される識別フラグ  $S_{vs}$ と、雑音分類回路 2 0 3 0から出力される分類フラグ  $S_{nz}$ とを入力し、前記識別フラグの値と分類フラグの値とに応じてスイッチを切り替えることで、前記第2のゲイン  $\hat{g}_2^{(m)}(n)$ を、 $S_{vs} = 0$ かつ  $S_{nz} = 0$ のときは第4のフィルタ 2 2 5 0へ、 $S_{vs} = 0$ かつ  $S_{nz} = 1$ のときは第5のフィルタ 2 2 6 0へ、 $S_{vs} = 1$ のときは第6のフィルタ 2 2 7 0へ出力する。第4のフィルタ 2 2 5 0は、第2の切替回路 2 2 1 0から出力される第2のゲイン  $\hat{g}_2^{(m)}(n)$ を入力し、線形フィルタまたは非線型フィルタを用いて平滑化し、これを第1の平滑化ゲイン  $\bar{g}_{2,1}^{(m)}(n)$ とし、第2のゲイン回路 1 1 3 0へ出力する。ここでは、次式で表されるフィルタを用いる。

$$\bar{g}_{2,1}^{(m)}(n) = \gamma_2 \cdot \bar{g}_{2,1}^{(m-1)}(n) + (1 - \gamma_2) \cdot \hat{g}_2^{(m)}(n)$$

ただし、 $\bar{g}_{2,1}^{(0)}(n) = \bar{g}_{2,1}^{(N_p)}(n-1)$ とする。また、 $\gamma_2 = 0.9$ とする。第5のフィルタ 2 2 6 0は、第2の切替回路 2 2 1 0から出力される第2のゲイン  $\hat{g}_2^{(m)}(n)$ を入力し、線形フィルタまたは非線型フィルタを用いて平滑化し、これを第2の平滑化ゲイン  $\bar{g}_{2,2}^{(m)}(n)$ とし、第2のゲイン回路 1 1 3 0へ出力する。ここでは、次式で表されるフィルタを用いる。

$$\bar{g}_{2,2}^{(m)}(n) = \gamma_2 \cdot \bar{g}_{2,2}^{(m-1)}(n) + (1 - \gamma_2) \cdot \hat{g}_2^{(m)}(n)$$

ただし、 $\bar{g}_{2,2}^{(0)}(n) = \bar{g}_{2,2}^{(N_p)}(n-1)$ とする。



【 0 0 2 8 】

また、 $\gamma_2 = 0.9$ とする。第6のフィルタ2270は、第2の切替回路2210から出力される第2のゲイン $\hat{g}_2^{(m)}(n)$ を入力し、線形フィルタまたは非線型フィルタを用いて平滑化し、これを第3の平滑化ゲイン $\bar{g}_{2,3}^{(m)}(n)$ とし、第2のゲイン回路1130へ出力する。ここでは、 $\bar{g}_{2,3}^{(m)}(n) = \hat{g}_2^{(m)}(n)$ とする。

第1のゲイン復号回路2220は、複数個のゲインが格納されたテーブルを複数セット備えている。第1のゲイン復号回路2220は、符号入力回路1010から出力される第3のゲインに対応するインデックスを入力し、音声モード復号回路2050から出力される音声モード $S_{mode}$ を入力し、フレームパワー復号回路2040から出力されるフレームパワー $\hat{E}_{rms}$ を入力し、線形予測係数変換回路1030から出力される第 $n$ フレーム第 $m$ サブフレームの線形予測係数 $\hat{\alpha}_j^{(m)}(n)$ 、 $j=1, \Lambda, N_p$ を入力し、ピッチ信号復号回路1210から出力されるピッチベクトル $c_{ac}(i)$ 、 $i=0, \Lambda, L_{sfr}$ を入力する。まず、線形予測係数 $\hat{\alpha}_j^{(m)}(n)$ から $k$ パラメータ $k_j^{(m)}(n)$ 、 $j=1, \Lambda, N_p$ （簡略化して $k_j$ と表す）を計算する。この計算は周知の方法、例えば、L. R. Rabiner らによる「Digital Processing of Speech Signals」（Prentice-Hall, 1978）（文献4）の第8.3.2節を参照できる。次に、推定残差パワー $\tilde{E}_{res}$ を前記 $k_j$ を用いて次式により計算する。

$$\tilde{E}_{res} = \hat{E}_{rms} \sqrt{\prod_{j=1}^{N_p} (1 - k_j^2)}$$

前記インデックスに対応する第3のゲイン $\hat{\gamma}_{gac}$ を、前記音声モード $S_{mode}$ によって切り替えられた前記テーブルより読み出し、次式を用いて第1のゲイン $\hat{g}_{ac}$ を計算する。

$$\hat{g}_{ac} = \hat{\gamma}_{gac} \frac{\tilde{E}_{res}}{\sqrt{\sum_{i=0}^{L_{sfr}-1} c_{ac}^2(i)}}$$

第1のゲイン復号回路2220は、第1のゲイン $\hat{g}_{ac}$ を第1のゲイン回路1230へ出力する。第2のゲイン復号回路2120は、複数個のゲインが格納されたテーブルを複数セット備えている。

【0029】

第2のゲイン復号回路2120は、符号入力回路1010から出力される第4のゲインに対応するインデックスを入力し、音声モード復号回路2050から出力される音声モード $S_{mode}$ を入力し、フレームパワー復号回路2040から出力されるフレームパワー $\hat{E}_{rms}$ を入力し、線形予測係数変換回路1030から出力される第 $n$ フレーム第 $m$ サブフレームの線形予測係数 $\hat{\alpha}_j^{(m)}(n)$ ,  $j=1, \Lambda, N_p$ を入力し、音源信号復号回路1110から出力される音源ベクトル $c_{ec}(i)$ ,  $i=0, \Lambda, L_{sf}$ を入力する。まず、線形予測係数 $\hat{\alpha}_j^{(m)}(n)$ から $k$ パラメーター $k_j^{(m)}(n)$ ,  $j=1, \Lambda, N_p$ （簡略化して $k_j$ と表す）を計算する。この計算には、第1のゲイン復号回路2220で述べた周知の方法を同様に用いることができる。次に、推定残差パワー $\tilde{E}_{res}$ を前記 $k_j$ を用いて次式により計算する。

$$\tilde{E}_{res} = \hat{E}_{rms} \sqrt{\prod_{j=1}^{N_p} (1 - k_j^2)}$$

前記インデックスに対応する第4のゲイン $\hat{\gamma}_{gec}$ を、前記音声モード $S_{mode}$ によって切り替えられた前記テーブルより読み出し、次式を用いて第2のゲイン $\hat{g}_{ec}$ を計算する。

$$\hat{g}_{ec} = \hat{\gamma}_{gec} \frac{\tilde{E}_{res}}{\sqrt{\sum_{i=0}^{L_{sf}-1} c_{ec}^2(i)}}$$

第2のゲイン復号回路2120は、第2のゲイン $\hat{g}_{ec}$ を第2の切替回路2210へ出力する。

図2は、本発明の音声信号符号化復号装置における音声信号復号装置の他の実施の形態を示すブロック図である。

本発明の音声信号復号装置は、第1の実施の形態において、フレームパワー復号回路2040を、パワー計算回路3040で置換し、音声モード復号回路2050を、音声モード決定回路3050で置換し、第1のゲイン復号回路2220を、第1のゲイン復号回路1220で置換し、第2のゲイン復号回路2120を、第2のゲイン復号回路1120で置換することで実現される。すなわち、フレームパワーと音声モードは、符号化器において符号化伝送されず、復号器において使用されるパラメータを用いて、前記フレームパワー（パワー）と前記音声モードを求める構成である。

【 0 0 3 0 】

ここで、第1のゲイン復号回路1220と、第2のゲイン復号回路1120は、図4の従来例で説明した各ブロックと同じであるので、ここでは説明を省略する。パワー計算回路3040は、合成フィルタ1040から出力される再生ベクトルを入力し、前記再生ベクトルの自乗和から、パワーを計算し、前記パワーを有音/無音識別回路2020へ出力する。ここでは、サブフレーム毎にパワーを計算するものとし、第 $m$ サブフレームにおける前記パワーの計算には、第 $m-1$ サブフレームにおいて前記合成フィルタ1040から出力された前記再生信号を用いる。前記再生信号を $s_{syn}(i)$ 、 $i=0, \Lambda, L_{sfr}$ とすると、前記パワー $E_{rms}$ は、例えば、平均自乗根 (Root Mean Square, RMS) により計算される。

$$E_{rms} = \sqrt{\sum_{i=0}^{L_{sfr}-1} s_{syn}^2(i)}$$

音声モード決定回路3050は、記憶回路1240に保持されている過去の励振ベクトル $e_{mem}(i)$ 、 $i=0, \Lambda, L_{mem}-1$ と、符号入力回路1010から出力されるインデックスとを入力する。前記インデックスは、遅延 $L_{pd}$ を指定する。ここで、 $L_{mem}$ は、 $L_{pd}$ の最大値により決定される定数である。第 $m$ サブフレームにおいて、前記過去の励振ベクトル $e_{mem}(i)$ と、前記遅延 $L_{pd}$ とから、ピッチ予測ゲイン $G_{emem}(m)$ 、 $m=1, \Lambda, N_{sfr}$ を計算する。

$$G_{emem}(m) = 10 \cdot \log_{10}(g_{emem}(m))$$

ここで、

$$g_{emem}(m) = \frac{1}{1 - \frac{E_c^2(m)}{E_{a1}(m)E_{a2}(m)}}$$

$$E_{a1}(m) = \sum_{i=0}^{L_{sfr}-1} e_{mem}^2(i)$$

$$E_{a2}(m) = \sum_{i=0}^{L_{sfr}-1} e_{mem}^2(i - L_{pd})$$

$$E_c(m) = \sum_{i=0}^{L_{sfr}-1} e_{mem}(i)e_{mem}(i - L_{pd})$$

である。

【 0 0 3 1 】

前記ピッチ予測ゲイン  $G_{emem}(m)$  あるいは、前記  $G_{emem}(m)$  の第  $n$  フレームにおけるフレーム内平均値  $\bar{G}_{emem}(n)$  に対し次の閾値処理を行なうことにより、音声モード  $S_{mode}$  を設定する。

if  $(\bar{G}_{emem}(n) \geq 3.5)$  then  $S_{mode} = 2$   
else  $S_{mode} = 0$

音声モード決定回路 3 0 5 0 は、前記音声モード  $S_{mode}$  を、有音/無音識別回路 2 0 2 0 へ出力する。

図 3 は、本発明の音声信号符号化復号装置における音声信号符号化装置の一実施の形態を示すブロック図である。

本発明の音声信号符号化装置は、図 5 の従来例において、フレームパワー計算回路 5 5 4 0 および音声モード決定回路 5 5 5 0 を付加し、第 1 のゲイン生成回路 5 2 2 0 および第 2 のゲイン生成回路 5 1 2 0 により、第 1 のゲイン生成回路 6 2 2 0 および第 2 のゲイン生成回路 6 1 2 0 を置換し、符号出力回路 5 0 1 0 により、符号出力回路 6 0 1 0 を置換することで実現される。なお、第 1 のゲイン生成回路 5 2 2 0、第 2 のゲイン生成回路 5 1 2 0、加算器 1 0 5 0 および記憶回路 1 2 4 0 は、図 5 の従来例で説明した各ブロックと同じであるので、ここでは説明を省略する。

フレームパワー計算回路 5 5 4 0 は、複数のフレームエネルギーが格納されたテーブルを備えている。入力端子 3 0 から入力ベクトルを入力し、前記入力ベクトルの平均自乗根 (Root Mean Square, RMS) を計算し、この RMS を前記テーブルを用いて量子化し、量子化フレームパワー  $\hat{E}_{rms}$  を得る。前記入力ベクトルを  $s_i(i)$ ,  $i = 0, \Lambda, L_{eff}$  とすると、パワー  $E_{rms}$  は、次式により計算される。

$$E_{rms} = \sqrt{\sum_{i=0}^{L_{eff}-1} s_i^2(i)}$$

前記量子化フレームパワー  $\hat{E}_{rms}$  を第 1 のゲイン生成回路 5 2 2 0 と第 2 のゲイン生成回路 5 1 2 0 とへ出力し、この  $\hat{E}_{rms}$  に対応するインデックスを符号出力回路 5 0 1 0 へ出力する。音声モード決定回路 5 5 5 0 は、重み付けフィルタ 5 0 5 0 から出力される重み付け入力ベクトルを入力する。

【 0 0 3 2 】

音声モード  $S_{mode}$  は、前記重み付け入力ベクトルを用いて求められる開ループ  
ピッチ予測ゲイン  $G_{op}(m)$  のフレーム内平均値  $\bar{G}_{op}(n)$  を閾値処理することで決  
定される。ここで、 $n$  はフレーム番号、 $m$  はサブフレーム番号を表す。第  $m$  サ  
ブフレームにおいて、重み付け入力ベクトル  $s_{wi}(i)$  と、遅延  $L_{tmp}$  とから、次の二  
式を計算し、 $E_{sctmp}^2(m)/E_{sa2tmp}$  を最大にする、前記  $L_{tmp}$  を求め、 $L_{op}$  とする。

$$E_{sctmp}(m) = \sum_{i=0}^{L_{sp}-1} s_{wi}(i) s_{wi}(i - L_{tmp})$$

$$E_{sa2tmp}(m) = \sum_{i=0}^{L_{sp}-1} s_{wi}^2(i - L_{tmp})$$

次に、重み付け入力ベクトル  $s_{wi}(i)$  と、遅延  $L_{op}$  とから、ピッチ予測ゲイン  
 $G_{op}(m)$ ,  $m = 1, \Lambda, N_{sfr}$  を計算する。

$$G_{op}(m) = 10 \cdot \log_{10}(g_{op}(m))$$

ここで、

$$g_{op}(m) = \frac{1}{1 - \frac{E_x^2(m)}{E_{sa1}(m)E_{sa2}(m)}}$$

$$E_{sa1}(m) = \sum_{i=0}^{L_{sp}-1} s_{wi}^2(i)$$

$$E_{sa2}(m) = \sum_{i=0}^{L_{sp}-1} s_{wi}^2(i - L_{op})$$

$$E_x(m) = \sum_{i=0}^{L_{sp}-1} s_{wi}(i) s_{wi}(i - L_{op})$$

である。前記ピッチ予測ゲイン  $G_{op}(m)$  あるいは、前記  $G_{op}(m)$  の第  $n$  フレームに  
おけるフレーム内平均値  $\bar{G}_{op}(n)$  に対し次の閾値処理を行なうことにより、音声  
モード  $S_{mode}$  を設定する。。

$$\text{if } (\bar{G}_{op}(n) \geq 3.5) \text{ then } S_{mode} = 2$$

$$\text{else } S_{mode} = 0$$

## 【0033】

音声モードの決定方法に関しては、K. Ozawa らによる「M-LCELP Speech Coding at 4 kb/s with Multi-Mode and Multi-Codebook」(IEICE Trans. On Commun., Vol. E77-B, No. 9, pp.1114-1121, 1994) (文献3)を参照できる。音声モード決定回路2040は、前記音声モード $S_{mode}$ を、第1のゲイン生成回路5220と第2のゲイン生成回路5120とへ出力し、前記音声モード $S_{mode}$ に対応するインデックスを符号出力回路5010へ出力する。ピッチ信号生成回路5210、音源信号生成回路5110、第1のゲイン生成回路5220および第2のゲイン生成回路5120は、各々、最小化回路5070から出力されるインデックスを順次入力する。これら、ピッチ信号生成回路5210、音源信号生成回路5110、第1のゲイン生成回路5220および第2のゲイン生成回路5120は、各々、入出力に関する結線を除けば、図1のピッチ信号復号回路1210、音源信号復号回路1110、第1のゲイン復号回路2220および第2のゲイン復号回路2120と同じであるので、これら各ブロックの詳細な説明は省略する。符号出力回路5010は、LSP変換/量子化回路5520から出力される量子化LSPに対応するインデックスを入力し、フレームパワー計算回路5540から出力される量子化フレームパワーに対応するインデックスを入力し、音声モード決定回路5550から出力される音声モードに対応するインデックスを入力し、最小化回路5070から出力される、音源ベクトル、遅延 $L_{pd}$ 、第1のゲインおよび第2のゲインの各々に対応するインデックスを入力し、各インデックスをビット系列の符号に変換し、出力端子40を介して出力する。

本発明の第4の実施例による、音声信号符号化復号装置における音声信号符号化装置の構成は、従来例による、音声信号符号化復号装置における音声信号符号化装置の構成と同じである。したがって、説明を省略する。

## 【0034】

## 【発明の効果】

本発明による第 1 の効果は、有音部で間欠的に含まれる短い無音部や有音部の一部において発生する通常の有音部とは異質の音を削減でき、その結果、有音部において不連続音を削減できることである。その理由は、この短い無音部において、時間変動（ばらつき）が小さい長時間平均した  $d_0(m)$  を用いるためである。また、有音部と無音部との識別を行い、有音部では、平滑化係数を零とするためである。

本発明による第 2 の効果は、無音部における平滑化係数の急激な変化を低減し、その結果、無音部において不連続音が削減されたことである。その理由は、時間的な変動の少ない長時間平均した  $d_0(m)$  を用いて平滑化係数を定めるためである。

本発明による第 3 の効果は、背景雑音の種類に応じて平滑化処理を選択できることであり、その結果、復号音質を改善できることである。その理由は、無音部において、複数個用意した処理方法を入力信号の特性に合わせて選択的に用いて復号パラメータの平滑化を行うからである。

## 【0035】

## 【図面の簡単な説明】

## 【図 1】

本発明の音声信号復号装置の第 1 の実施の形態を示すブロック図である。

## 【図 2】

本発明の音声信号復号装置の第 2 の実施の形態を示すブロック図である。。

## 【図 3】

本発明の音声信号符号化装置一実施の形態を示すブロック図である。

## 【図 4】

従来の音声信号復号装置を説明するブロック図である。

## 【図 5】

従来の音声信号符号化装置を説明するブロック図である。

## 【符号の説明】

10,30 入力端子

20,40 出力端子

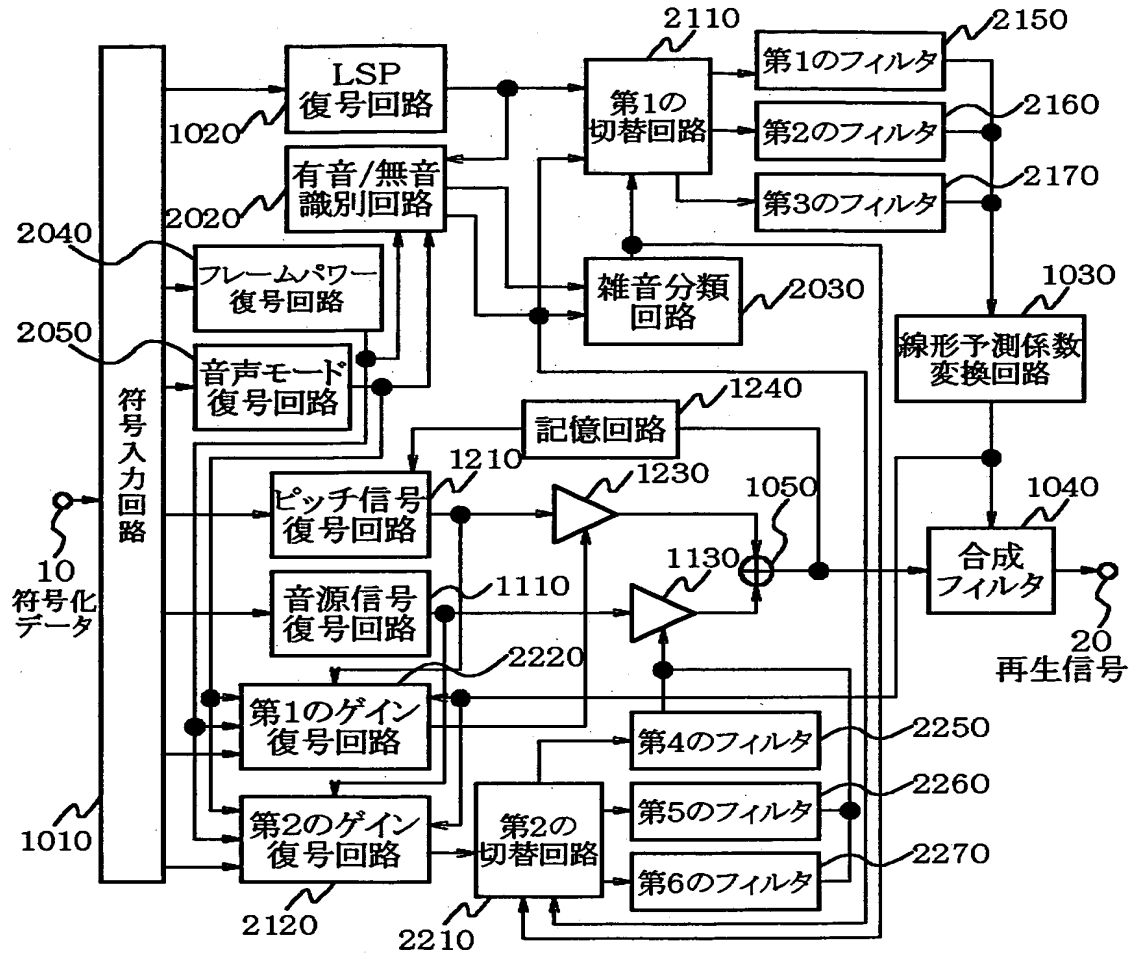
1010,2010 符号入力回路  
1020 LSP復号回路  
1030,5030 線形予測係数変換回路  
1040 合成フィルタ  
1050 加算器  
1110 音源信号復号回路  
1210 ピッチ信号復号回路  
1120,2120 第2のゲイン復号回路  
1220,2220 第1のゲイン復号回路  
1130 第2のゲイン回路  
1230 第1のゲイン回路  
1240 記憶回路  
1310 平滑化係数計算回路  
1320 平滑化回路  
2020 有音/無音識別回路  
2030 雑音分類回路  
2040 フレームパワー復号回路  
2050 音声モード復号回路  
2110 第1の切替回路  
2150 第1のフィルタ  
2160 第2のフィルタ  
2170 第3のフィルタ  
2210 第2の切替回路  
2250 第4のフィルタ  
2260 第5のフィルタ  
2270 第6のフィルタ  
3040 フレームパワー計算回路  
3050 音声モード決定回路  
5510 線形予測係数計算回路



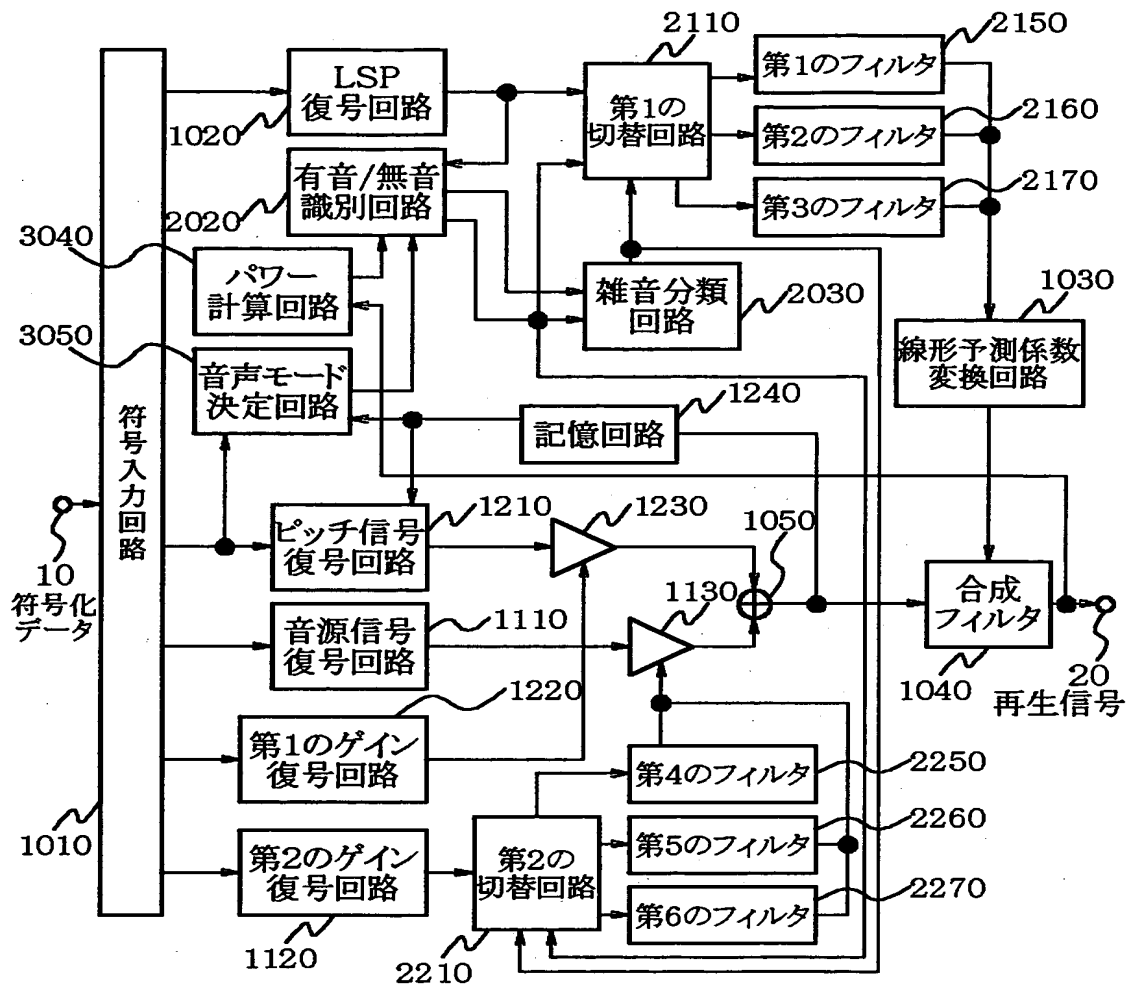
5520	LSP変換/量子化回路
5040	重み付け合成フィルタ
5050	重み付けフィルタ
5060	差分器
5070	最小化回路
5540	フレームパワー計算回路
5550	音声モード決定回路
5210	ピッチ信号生成回路
5110	音源信号生成回路
5220,6220	第1のゲイン生成回路
5120,6120	第2のゲイン生成回路
5010,6010	符号出力回路

【書類名】 図面

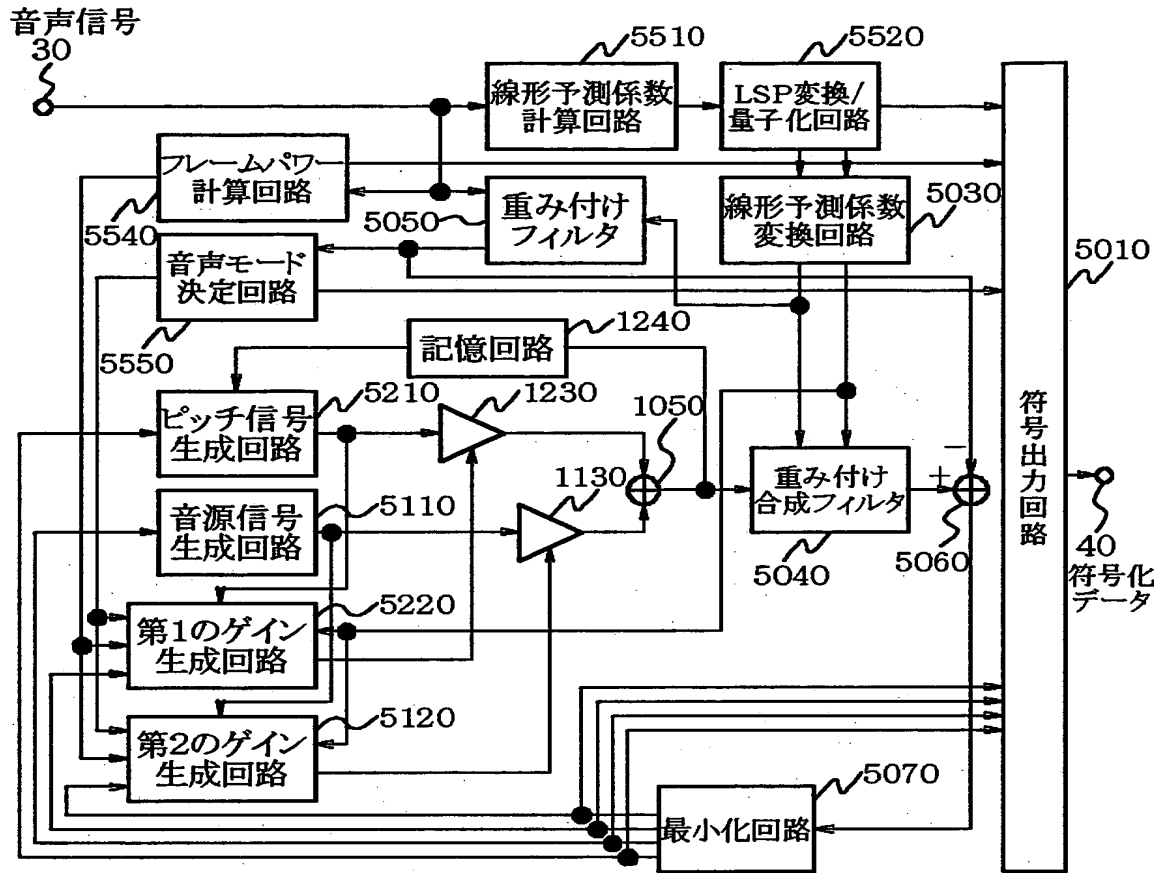
【図1】



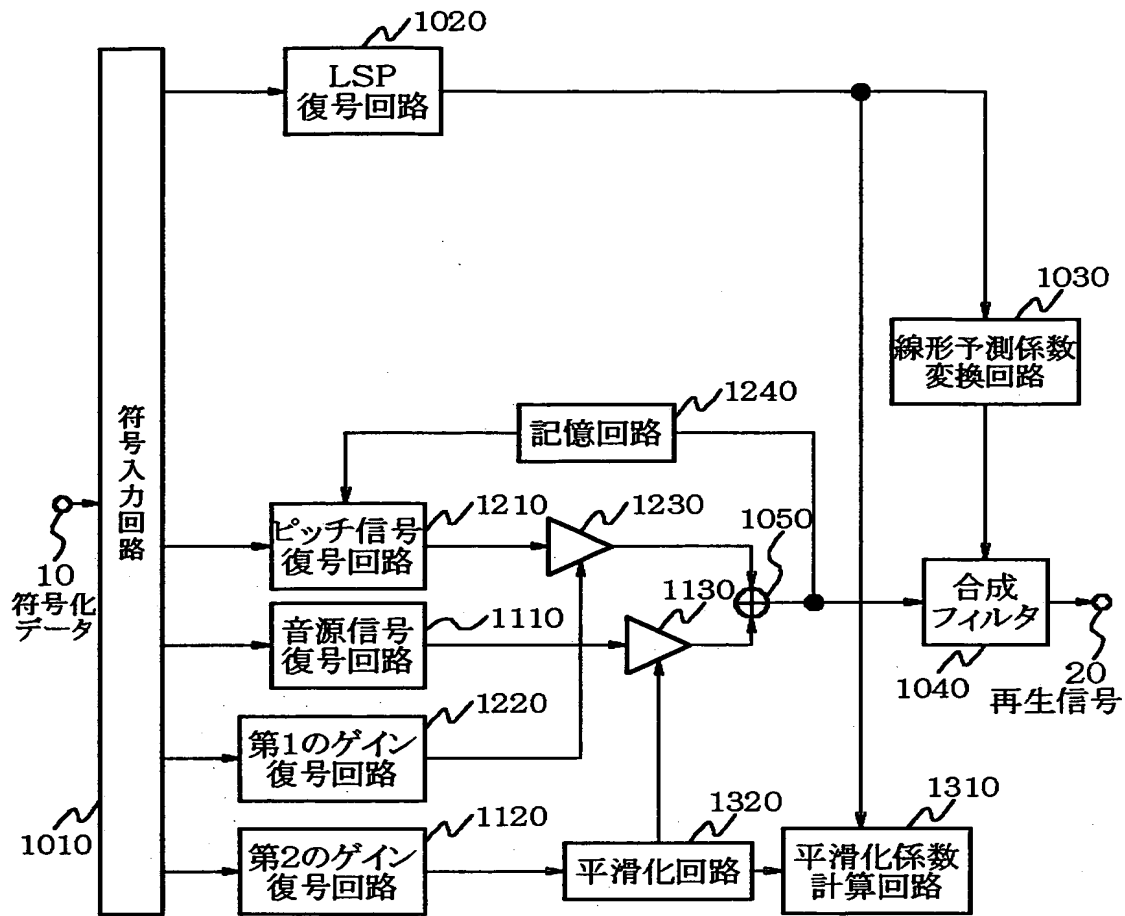
【図 2】



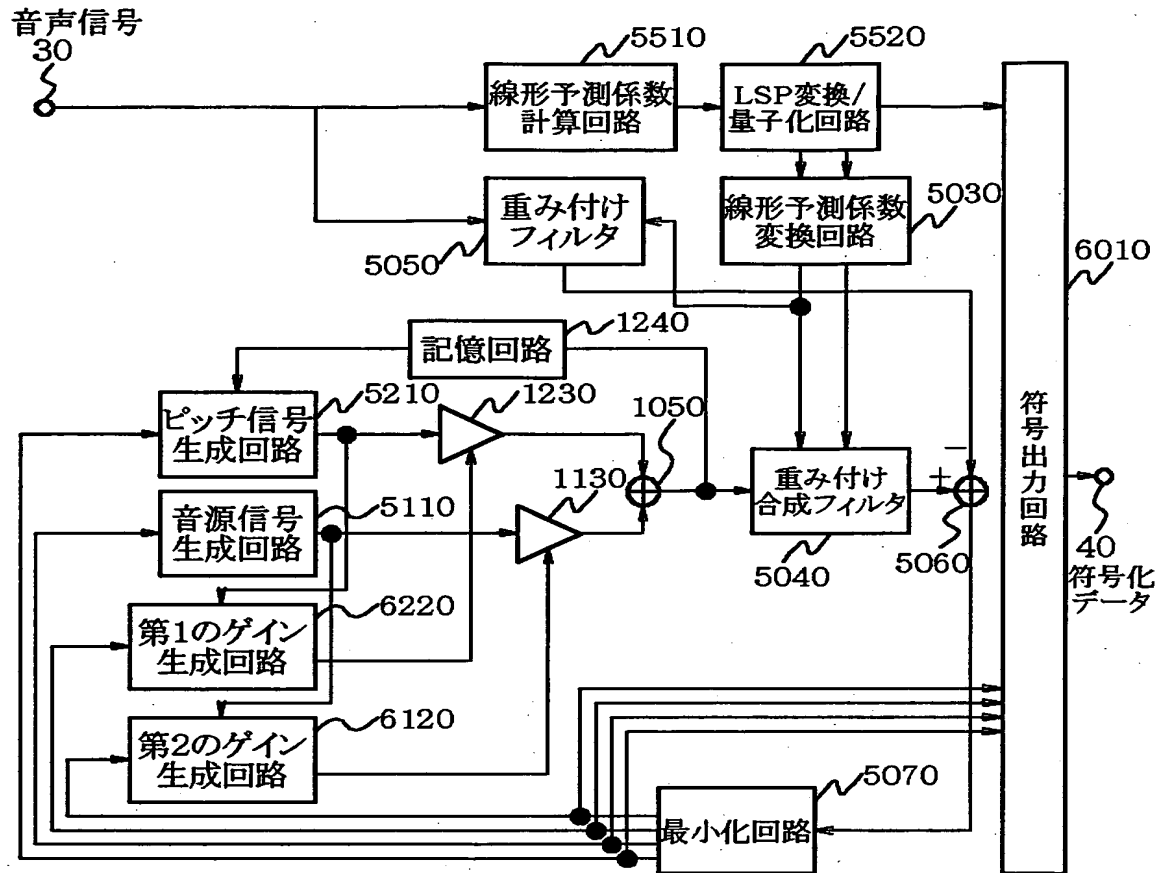
【図 3】



【図 4】



【図 5】



【書類名】 要約書

【要約】

【課題】 音源信号とゲインとフィルタ係数とを復号し、前記音源信号に前記ゲインを乗じて得られる励振信号により前記フィルタ係数で構成するフィルタを駆動することによって音声信号を生成する音声信号復号装置において、背景雑音音声に対する再生音声品質を改善する。

【解決手段】 有音/無音識別回路（図 1 の 2 0 2 0）は、LSP と LSP の長時間平均とからスペクトルパラメータの変動量を計算し、前記変動量の長時間平均と、フレームパワーと、ピッチ予測ゲインとから有音/無音の識別を行う。雑音分類回路（図 1 の 2 0 3 0）は、前記変動量の長時間平均を無音部においてさらに長時間平均し、これを用いて雑音の分類を行う。第 1 の切替回路と第 2 の切替回路（図 1 の 2 1 1 0、2 2 1 0）は、前記有音/無音の識別および前記雑音の分類に基づいて復号パラメータに適用するフィルタを切り替える。

【選択図】 図 1

認定・付加情報

特許出願の番号	平成11年 特許願 第214292号
受付番号	59900725664
書類名	特許願
担当官	第七担当上席 0096
作成日	平成11年 7月30日

<認定情報・付加情報>

【提出日】	平成11年 7月28日
-------	-------------